PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (**http://bmjopen.bmj.com/site/about/resources/checklist.pdf**) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Is social deprivation an independent predictor of outcome following cardiac surgery? An analysis of 240,221 patients from a national registry
AUTHORS	Barnard, James; Grant, Stuart; Hickey, Graeme; Bridgewater, Ben

VERSION 1 - REVIEW

REVIEWER	Chris Gale University of Leeds, Division of Epidemiology and Biostatistics
REVIEW RETURNED	09-Sep-2014

GENERAL COMMENTS	Barnard and colleges undertake a retrospective analysis of a national cohort of nearly a quarter of a million patients who received cardiothoracic surgery between 2003 and 2013. The authors suggest that socio-economic status (SES), quantified using the Index of Multiple Deprivation composite score, is associated with both early and late all-cause mortality. The authors note that SES is not routinely accounted for when considering variation in mortality rates from cardiothoracic surgery, and suggest that because IMD predicts outcome, this Index could be considered in future provider performance analyses. In general, this is a well written and considered manuscript from an established group using recognised data. Clearly, SES is an important factor associated with cardiovascular outcomes. There are, however, a number of points which should be considered. 1) Abstract results. Provide relative risks as an estimate of adjusted impact of IMD on in-hospital death and survival. 2) Abstract results. Please define PLOS 3) Abstract results. Reconsider including sentence about PLOS as this is not the focus of the study - also see comments below about 'significance'. Happy for authors to keep in, but please be mindful of your interpretation of the results 4) Abstract conclusion. I would reconsider the term 'predictor' and perhaps rephrase to suggest 'associated with' 5) Analyses. I think the authors should consider a path diagram (which doesn't need to be published) that may help refine the analyses. For example, consider giving the relationships between all important factors: surgery type and outcome, EuroSCORE and outcome, IMD and outcome, geographical region and outcome. For the manuscript, please test and state the interaction term (which I suspect will be a significant) between IMD, EuroSCORE and outcomes. Also test IMD*geographical region on outcomes. 6) Please look up and reference Impact of missing data on standardised mortality ratios for acute myocardial infarction:

	evidence from the Myocardial Ischaemia National Audit Project (MINAP) 2004-7. Gale CP, Cattle BA, Moore J, Dawe H, Greenwood DC, West RM.Heart. 2011 Dec;97(23):1926-31. doi: 10.1136/hrt.2010.204883. Epub 2011 Jan 12. which considered IMD on CV outcomes and may give some insight into the usefulness of age, sex and IMD (often used in administrative database analyses) for quantifying provider performance. 7) Did the authors have access the components of IMD – what factors impacted on outcomes? 8) Access to surgery may be IMD dependant, which in itself introduces bias to the analyses. 9) As the authors correctly mention that IMD is a geographical aggregate please mention bias (ecological fallacy, etc). Also, survival analyses assume constancy – especially that patents do not relocate to another IMD area – 10) Methods. First paragraph – removed duplicate wording ('clean') in a single sentence 11) Missing data. The authors use case deletion and in other instances default imputation. I know that they will understand that this is not ideal (whilst allowing more completed cases it shrinks variance estimates). Other (now off the shelf) imputation strategies may be worth considering, and if not sensitivity analyses to quantify the impact (or not) of introduced bias undertaken and added to appendices / e-tables. See paragraph two and three of the methods. 12) Please define in methods how IMD categories where defined. Equal sizes or equal cut-offs, I suspect the latter. 13) Methods, outcomes. Consider the impact of IMD on LOS. The authors use only post op LOS. Undoubtedly there is a complex relationship between LOS and in-hospital death. This may well be affected by IMD. Therefore please consider reporting total LOS as well as its components (pre and post op). I suspect increasing IMD will be associated with longer total, pre and post op LOS. In this respect, I wonder whether in-hospital death should be weighted by LOS (as well as giving unadjusted and risk ratios)? Please test the interaction between LOS*IMD and outcomes – this should support the authors statements and initial inference testing. 14) Please define 'mid-term survival'. In the results / discussion the authors mention long-term survival and there should be consistency throughout. 10 year is long-term! 15) Please write in the past tense 'was' not 'is' 16) Data 'are' pleural 17) Please test (and give results of the testing) the Cox proportional assumptions. I have a feeling they will be violated. If so, consider, stratifying by adjustment factor, a more parsimonious model (as sensitivity analysis), and / or land mark analyses. 18) Three is evidence of accelerated – interestingly, more so for CABG + Valve than isolated CABG. Do the/different link functions need to be tested and used? 19) Please add components of the EuroSCORE into the model as separate factors / covariates as a sensitivity analysis. I assume there was no co-linearity? 20) I think the relationship between outcomes and obesity can be de-emphasised, but not removed. It detracts from the main message and has been the subject of other papers in this area. 21) Results. IMD categories – state them in the methods not results, gives results in results! 22) Please be very careful about stating 'significant'. Clearly you mean statistical significance (though I would argue that testing the null is not appropriate in such a large database – as you have shown

	p values tend to be very small! Moreover, are the (no-different, but statistically significant) LOS days by IMD clinically important? Perhaps give LOS in the table to 1 decimal point? Or maybe superimposed kernel plots by IMD category?? 23) Results, survival. State the total number of years follow-up in the analytical cohort – will be massive! 24) Results, survival. In the text, please give (HRs and) interpretation for all (not some) of the IMD, surgery type, outcome permutations. 25) Table 1 could be converted to a Forest plots, also giving the RRs and numbers of cases included. 26) Figure 4 is necessary, but may need to be modified in light of testing the assumptions (landmark KM plot etc)

REVIEWER	Rod Taylor University of Exeter Peninsula Medical School Exeter Cornwall United Kingdom
REVIEW RETURNED	13-Oct-2014

GENERAL COMMENTS	The paper seeks investigate in a national UK data set whether there are geographical variations in the social deprivation of patients undergoing cardiac surgery and identify whether social deprivation is an independent predictor of outcomes. The paper is well presented and appears to have been appropriately conducted. I have only a small issues - discuss potential limitations of EuroSCORE and ability to adjust for confounding across IMD strata - the analytic approach to PLOS data seems over-simplistic - can the authors need to explain why a hierarchical regression model was not used (as for mortality outcomes) - I found it bothersome that IMD was not found to be consistently associated with outcomes in this study. Whilst the authors provide some comment on as this as the centre of the authors hypothesis and the aim of this paper, further explanation for the possible explanations for this variance would be reassuring.

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

Comment 1: Abstract results. Provide relative risks as an estimate of adjusted impact of IMD on in-hospital death and survival.

Response 1: We agree with this comment and made the change as requested.

Change 1: Included hazard ratios and odds ratios in the Abstract (page 3, line 76-79).

Comment 2: Abstract results. Please define PLOS

Response 2: We thank the reviewer for their comment and have now defined PLOS in the abstract.

Change 2: PLOS defined in the abstract

Comment 3: Abstract results. Reconsider including sentence about PLOS as this is not the focus of the study - also see comments below about 'significance'. Happy for authors to keep in, but please be mindful of your interpretation of the results

Response 3: We have changed the PLOS analysis and revised this sentence as appropriate.

Change 3: We feel that PLOS is an important secondary outcome and have therefore kept it in the abstract.

Comment 4: Abstract conclusion. I would reconsider the term 'predictor' and perhaps rephrase to suggest 'associated with'

Response 4: We agree this suggestion and have amended the sentence as appropriate.

Change 4: We have rephrased this sentence of the abstract (page 4, line 83)

Comment 5: Analyses. I think the authors should consider a path diagram (which doesn't need to be published) that may help refine the analyses. For example, consider giving the relationships between all important factors: surgery type and outcome, EuroSCORE and outcome, IMD and outcome, geographical region and outcome. For the manuscript, please test and state the interaction term (which I suspect will be a significant) between IMD, EuroSCORE and outcomes. Also test IMD*geographical region on outcomes.

Response 5: We believe that the development of a path diagram at this stage would not add further information to the study. However, we note this for future studies, and will make them available as required. It is not possible to test IMD*geographical region because the baseline hazards function (in the case of Cox regression) is stratified by hospitals, which are nested within each region. Similarly, the mixed-effects logistic regression model features random-effects terms for hospital. Owing to no first-order interaction terms being found in Pagano et al. (2009), we did not test for them.

Change 5: None.

Comment 6: Please look up and reference Impact of missing data on standardised mortality ratios for acute myocardial infarction: evidence from the Myocardial Ischaemia National Audit Project (MINAP) 2004-7. Gale CP, Cattle BA, Moore J, Dawe H, Greenwood DC, West RM.Heart. 2011 Dec;97(23):1926-31. doi: 10.1136/hrt.2010.204883. Epub 2011 Jan 12. which considered IMD on CV outcomes and may give some insight into the usefulness of age, sex and IMD (often used in administrative database analyses) for quantifying provider performance.

Response 6: Although we found the suggested manuscript to be of interest but as their focus was on the impact of missing data (including IMD) on standardised mortality ratios we did not think it was related closely enough to the overall message of our manuscript. We could not identify an obvious place to reference it without adding significantly to the discussion so we have therefore made no change.

Change 6: None

Comment 7: Did the authors have access the components of IMD – what factors impacted on outcomes?

Response 7: We did not have access to the individual components; the IMD is calculated off-site and linked to the National Adult Cardiac Surgery Audit registry.

Change 7: We have noted this point in the manuscript (page 6, lines 141-142).

Comment 8: Access to surgery may be IMD dependant, which in itself introduces bias to the analyses.

Response 8: It is possible that social deprivation is associated with inequalities in access to healthcare. We have commented on this in the limitations section of the text.

Change 8: Page 10, lines 301-302

Comment 9: As the authors correctly mention that IMD is a geographical aggregate please mention bias (ecological fallacy, etc.). Also, survival analyses assume constancy – especially that patents do not relocate to another IMD area.

Response 9: We have made reference in the strength and weaknesses of the study section to the limitations of using the IMD score. Specifically we have highlighted the fact that IMD score is based on geographical areas containing over 1000 people and that it is not applicable to individual patients.

Change 9: Page 10, Lines 295-297

Comment 10: Methods. First paragraph – removed duplicate wording ('clean') in a single sentence

Response 10: We agree with this comment.

Change 10: We have deleted word from sentence.

Comment 11: Missing data. The authors use case deletion and in other instances default imputation. I know that they will understand that this is not ideal (whilst allowing more completed cases it shrinks variance estimates). Other (now off the shelf) imputation strategies may be worth considering, and if not sensitivity analyses to quantify the impact (or not) of introduced bias undertaken and added to appendices / e-tables. See paragraph two and three of the methods.

Response 11: Records where the IMD score was missing or equal to zero were deleted. A missing IMD score can be due either: 1) a resident from outside of England, 2) a patient without a fixed address, 3) or missing administrative data. We inspected the data and identified that the greatest absolute number of records with missing or zero IMD score was a Merseyside hospital. This is consistent with our hypothesis that patients who are resident in Wales (and therefore do not have a comparable IMD score) might have travelled to this unit. As we do not want to bias the study with non-English residents, we opted to exclude this data, perhaps at the cost of excluding true missing data. Records corresponding to missing outcome data were also deleted since imputation would be a self-fulfilling prophecy. We acknowledge that modal-imputation for the adjustment variables is, in general, suboptimal relative to multiple imputation algorithms. A decision was taken to use this method based mainly on the fact that there were few missing data overall: all variables were <2% missing with the exception of only creatinine (5.0%), pulmonary hypertension (7.0%) and active endocarditis (2.8%). Furthermore, our experience with the completeness of missing EuroSCORE variables as part of national monitoring programmes is that the assumption is robust for risk-adjustment.

Change 11: Firstly, we have included a summary of the missing data for the purposes of transparency (page 7, line 200-202). Secondly, we have updated the study limitations to reflect this particular issue (page 10, line 295-297).

Comment 12: Please define in methods how IMD categories where defined. Equal sizes or equal cut-offs, I suspect the latter.

Response 12: We can confirm that the groups were 'quintile' groups, defined here to mean that approximately 20% of patients lie in each group; i.e. equal sizes. The classification was made on the dataset as a whole, and not repeated within subgroups. To avoid any misinterpretation, we have amended the text in the manuscript.

Change 12: Provided additional information on determination of quintile groups (page 7, lines 203-205).

Comment 13: Methods, outcomes. Consider the impact of IMD on LOS. The authors use only post op LOS. Undoubtedly there is a complex relationship between LOS and in-hospital death. This may well be affected by IMD. Therefore please consider reporting total LOS as well as its components (pre and post op). I suspect increasing IMD will be associated with longer total, pre and post op LOS. In this respect, I wonder whether in-hospital death should be weighted by LOS (as well as giving unadjusted and risk ratios)? Please test the interaction between LOS*IMD and outcomes – this should support the authors statements and initial inference testing.

Response 13: Total length of stay (LOS) is comprised of preoperative LOS and postoperative LOS. We believe that the preoperative LOS is dependent on organisational structure and models of care, e.g. same-day admission, keeping patients in surgical hospital after angiogram for urgent patients, neither of which are not believed to be associated directly with patient characteristics. Postoperative LOS is not believed to be strongly associated by these issues, but is believed to be a proxy measurement for postoperative complications. For this reason we restrict our analysis to postoperative LOS only.

Change 13: None.

Comment 14: Please define 'mid-term survival'. In the results / discussion the authors mention long-term survival and there should be consistency throughout. 10 year is long-term!

Response 14: Whilst we agree that 10-years is a long-term period, owing to the staggered entry of patients into the study over the 10-year window not all patients have 10-years follow-up. We have therefore used 'mid-term' rather than 'long-term' to emphasise that average follow-up was 4.8-years, not 10-years.

Change 14: None

Comment 15: Please write in the past tense 'was' not 'is'

Response 15: We agree with this comment and have made this correction throughout the manuscript.

Comment 16: Data 'are' pleural

Response 16: We acknowledge this error.

Change 16: We have corrected the nouns as appropriate.

Comment 17: Please test (and give results of the testing) the Cox proportional assumptions. I have a feeling they will be violated. If so, consider, stratifying by adjustment factor, a more parsimonious model (as sensitivity analysis), and / or land mark analyses.

Response 17: We have tested the assumption of proportional hazards in the Cox PH model. For IMD quintile groups we used complementary log-log plots of the Kaplan-Meier estimator. For the final Cox regression model, we applied the Grambsh-Therneau test to each explanatory variable and corresponding Schoenfeld residuals. Owing to the large sample sizes, the test was sensitive to even slight departures from the PH assumption when using the original (parsimonious) transformation of logistic EuroSCORE and BMI in the model. Therefore we have updated our models and used more flexible restricted cubic spline functions to model these variables, which has subsequently alleviated the original violation of the assumption. We have also applied the reviewers advice and further stratified our model (baseline hazards function) on further factors, namely smoking status, in cases

where violations remain. We note that the inferences remain broadly consistent with the original analysis.

Change 17: We have updated the Cox regression model to model BMI and log-transformed EuroSCORE using restricted cubic spline functions. In some models we have stratified on smoking history, as appropriate. Proportional hazards have been tested using an appropriate statistical test and no violations were found (with one minor exception; see text).

Comment 18: Three is evidence of accelerated – interestingly, more so for CABG + Valve than isolated CABG. Do the/different link functions need to be tested and used?

Response 18: It is not clear to us what is meant by this comment, i.e. “evidence of accelerated”. If the reviewer is referring to accelerated failure time (AFT) models, then we feel that a semi-parametric Cox PH model is perfectly adequate in light of demonstrating the PH assumption is reasonable.

Change 18: None.

Comment 19: Please add components of the EuroSCORE into the model as separate factors / covariates as a sensitivity analysis. I assume there was no co-linearity?

Response 19: As per Comment 2 of the Editorial Committee, we have now performed this sensitivity analysis on the ‘all cardiac surgery’ group for the survival analysis.

Change 19: See Response 2 / Change 2 from the Editorial Committee review.

Comment 20: I think the relationship between outcomes and obesity can be de-emphasised, but not removed. It detracts from the main message and has been the subject of other papers in this area.

Response 20: We have given this comment great thought. The reviewer is completely correct that the association between BMI and outcome is well established, and we do not wish for it to distract the reader. Therefore, we continue to adjust for BMI (now using spline functions), but only report the (adjusted-) effect sizes for IMD quintile.

Change 20: We have removed the effect sizes for BMI,

Comment 21: Results. IMD categories – state them in the methods not results, gives results in results!

Response 21: As the IMD categories are dependent on the final dataset (via the quantile function), which in turn depends on the execution of the model inclusion and exclusion criteria, we feel that they are part of the results for this study.

Change 21: None.

Comment 22: Please be very careful about stating ‘significant’. Clearly you mean statistical significance (though I would argue that testing the null is not appropriate in such a large database – as you have shown p values tend to be very small! Moreover, are the (no-different, but statistically significant) LOS days by IMD clinically important? Perhaps give LOS in the table to 1 decimal point? Or maybe superimposed kernel plots by IMD category??

Response 22: We have responded to the first point in our response for Editorial Committee 1. Mean LOS is already reported to 1 decimal place. Kernel plots are almost completely overlapping, hence they were considered to be uninformative. The PLOS analysis has also been changed.

Change 22: None.

Comment 23: Results, survival. State the total number of years follow-up in the analytical cohort – will be massive!

Response 23: This number is indeed large, and is now included.

Change 23: Added the total number of years of follow-up time (page 8, line 227).

Comment 24: Results, survival. In the text, please give (HRs and) interpretation for all (not some) of the IMD, surgery type, outcome permutations.

Response 24: We thank the reviewer for this comment. Due to limitations of the word count we were unable to report all hazard ratios in the text and now instead report all of these ratios in Figure 4.

Change 24: All hazard ratios are now reported in Figure 4.

Comment 25: Table 1 could be converted to a Forest plots, also giving the RRs and numbers of cases included.

Response 25: As suggested, we have replaced Tables 1 and 2 with forest plots for the effect sizes (odds ratio and hazard ratio, respectively), and have labelled them with the actual statistics for purposes of reference. We have limited the plots to the adjusted effect sizes for IMD quintile groups, which are the focus of the study, since EuroSCORE and BMI have already been shown to be associated with outcome (references are have been added to the study).

Change 25: Tables 1 and 2 have been replaced by Figures 2 and 4.

Comment 26: Figure 4 is necessary, but may need to be modified in light of testing the assumptions (landmark KM plot etc)

Response 26: We thank the reviewer for this comment. Kaplan-Meier curves are not dependent on any assumptions of proportionality, therefore we do not believe that they require modification.

Change 26: None.

Reviewer 2

Comment 1: Discuss potential limitations of EuroSCORE and ability to adjust for confounding across IMD strata.

Response 1: We thank the reviewer for the comment and have now identified the potential limitations of the logistic EuroSCORE in the discussion.

Change 1: Sentence added (page 10, lines 307-308)

Comment 2: The analytic approach to PLOS data seems over-simplistic - can the authors need to explain why a hierarchical regression model was not used (as for mortality outcomes)

Response 2: We thank the reviewer for this comment. We believe 'standard' methodologies to be severely lacking. For example regression models fitted to PLOS data often violate the assumptions of normally distributed errors, even with log-transformations. Moreover, they ignore the censoring due to in-hospital mortality. The use of time-to-event analyses assumes the censoring mechanism is independent of outcome, which is unlikely to hold. In order to keep the analysis concise, we avoided entering this area. However, we have now added a different analysis in response, namely an adjusted comparison of prolonged-PLOS (defined as 14 days or longer in hospital regardless of discharge status) by IMD quintile. This outcome is reported and routinely used by the United States Society of

Thoracic Surgeons, and therefore has a proven history for being relevant in the measurement of quality and outcomes following cardiac surgery.

Change 2: We have replaced the unadjusted non-parametric analysis of PLOS by a regression analysis on prolonged-PLOS. This analysis is described on page 7, lines 186-187 and the results given on page 9, line 255-258 with a corresponding figure (Figure 5).

Comment 3: I found it bothersome that IMD was not found to be consistently associated with outcomes in this study. Whilst the authors provide some comment on as this as the centre of the authors hypothesis and the aim of this paper, further explanation for the possible explanations for this variance would be reassuring.

Response 3: We thank the reviewers for this comment. We have now discussed in more detail the possible reasons why IMD was found to be more strongly associated with survival rather than in-hospital mortality.

Change 3: Page 10, lines 280-286.

We thank the reviewers for all of their comments and hope that the manuscript has been enhanced by the changes that have been made.

VERSION 2 – REVIEW

REVIEWER	Rod Taylor University of Exeter Peninsula Medical School Exeter Cornwall United Kingdom
REVIEW RETURNED	01-Feb-2015

GENERAL COMMENTS	The reviewers appear to have responded to the various issues raised in the first round of review

REVIEWER	Chris Gale University of Leeds, Division of Epidemiology and Biostatistics
REVIEW RETURNED	02-Feb-2015

GENERAL COMMENTS	The authors should be congratulated for taking the time and effort to revise the manuscript - which is now much better. Well done I have only a few minor comments, if these could be addressed. line 2 Title: place a colon and state the study design as per SROBE guidelines, eg a national cohort study STROBE: Title and abstract 1 (a) Indicate the study's design with a commonly used term in the title or the abstract YES Line 59 Comma before but Line 256 – what do the authors mean by prolonged (above the mean?)